

# SCORE Search Results Details for Application 09961086 and Search Result 20080917\_142911\_us-09-961-086a-1.rpr.

<a href="#">Score Home</a>	<a href="#">Retrieve Application</a>	<a href="#">SCORE System</a>	<a href="#">SCORE</a>	<a href="#">Comments /</a>
<a href="#">Page</a>	<a href="#">List</a>	<a href="#">Overview</a>	<a href="#">FAQ</a>	<a href="#">Suggestions</a>

This page gives you Search Results detail for the Application 09961086 and Search Result 20080917\_142911\_us-09-961-086a-1.rpr.

[Go Back to previous page](#)

GenCore version 6.2.1  
Copyright (c) 1993 - 2008 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2008, 22:00:18 ; Search time 41 Seconds  
(without alignments)  
1537.121 Million cell updates/sec

Title: US-09-961-086A-1  
Perfect score: 3352  
Sequence: 1 MSSSNVEVFIPVSQGNTNGF.....MIVIFLTIAYLKLLFLKKYS 655

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pirl:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	2849.5	85.0	656	2	JC7860	brain multidrug re	
2	835.5	24.9	1049	1	S19421	ATP-dependent perm	

3	812	24.2	687	1	FYFFW	white protein - fr
4	800.5	23.9	737	2	T46101	ABC transporter-li
5	774	23.1	646	2	C86441	probable ABC trans
6	767.5	22.9	687	2	D96553	hypothetical prote
7	746	22.3	649	2	A84509	probable ABC trans
8	745	22.2	725	2	T47652	ABC transporter-li
9	741.5	22.1	739	2	T45891	ABC transporter-li
10	739	22.0	678	2	H96552	hypothetical prote
11	726.5	21.7	708	2	T47650	ABC transporter-li
12	721	21.5	635	2	T08934	hypothetical prote
13	720.5	21.5	755	2	G84791	probable ABC trans
14	717.5	21.4	740	1	T02567	probable ATP-bindi
15	712.5	21.3	609	2	E96742	probable ABC trans
16	710	21.2	638	2	G02068	white homolog - hu
17	708	21.1	646	2	JC7777	ATP binding casset
18	707.5	21.1	547	2	T31543	hypothetical prote
19	707.5	21.1	720	2	T47648	ABC transporter-li
20	677	20.2	559	2	B88474	protein C05D10.3 [
21	676	20.2	725	2	C84423	probable ABC trans
22	669	20.0	659	2	E86313	hypothetical prote
23	664	19.8	608	2	T34391	hypothetical prote
24	660	19.7	662	2	T47649	ABC transporter-li
25	658.5	19.6	1294	2	S77690	probable membrane
26	657.5	19.6	1450	2	T45888	ABC transporter-li
27	656.5	19.6	590	2	B96573	protein F12M16.17
28	649.5	19.4	633	2	T19189	hypothetical prote
29	645.5	19.3	577	2	T04229	ABC-type transport
30	639	19.1	658	2	T31958	hypothetical prote
31	636.5	19.0	639	2	G88839	protein C10C6.5 [i
32	636.5	19.0	695	2	T21109	hypothetical prote
33	627.5	18.7	610	2	T19333	hypothetical prote
34	612.5	18.3	1501	2	S50992	SNQ2 protein - yea
35	612	18.3	1530	2	S52239	brefeldin a resist
36	603	18.0	1530	2	T52010	hypothetical prote
37	602	18.0	1511	2	A53151	pleiotropic drug r
38	594	17.7	1564	2	S55517	probable transport
39	593	17.7	1443	2	T02491	probable ABC trans
40	590	17.6	1469	2	H96622	probable ABC trans
41	585	17.5	1333	2	S63403	probable membrane
42	580.5	17.3	1420	2	T02644	ABC-type transport
43	580.5	17.3	1529	2	S69688	hypothetical prote
44	572	17.1	1413	2	G84790	probable ABC trans
45	564	16.8	1466	2	T30566	ATP-binding casset

## ALIGNMENTS

## RESULT 1

JC7860

brain multidrug resistance protein, BMDP - pig

C;Species: *Sus scrofa domestica* (domestic pig)

C;Date: 18-Nov-2002 #sequence\_revision 18-Nov-2002 #text\_change 09-Jul-2004

C;Accession: JC7860

R;Eisenblaetter, T.; Galla, H.J.

Biochem. Biophys. Res. Commun. 293, 1273-1278, 2002

A;Title: A new multidrug resistance protein at the blood-brain barrier.

A;Reference number: JC7860; MUID:22050127; PMID:12054514  
A;Accession: JC7860  
A;Molecule type: mRNA  
A;Residues: 1-656 <EIS>  
A;Cross-references: UNIPROT:Q8MIB3; UNIPARC:UPI0000087EC5; GB:AJ420927  
A;Experimental source: brain  
C;Comment: This protein, a new transport protein of the ATP-binding cassette (ABC) superfamily of transporters, expressed in porcine brain capillary endothelial cells, plays an importnat role in the exclusion of xenobiotics from the brain and participates in drug transport across the blood-brain barrier and therefore is considered as a efflux pump at the cerebral endothelium.  
C;Genetics:  
A;Gene: bmdp

Query Match 85.0%; Score 2849.5; DB 2; Length 656;  
Best Local Similarity 84.3%; Pred. No. 2.7e-192;  
Matches 553; Conservative 44; Mismatches 58; Indels 1; Gaps 1;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
		: :     :   :         :         :	
Db	1	MSSNSYQVSIPMSKRNTNGLPGSSSNELKTSAGGAVLSFHDICYRVKVKSGFLFCRKTV	60
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:	
Db	61	KEILTNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPHGLSGDVLINGAPRPANFKCN	120
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Db	121	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTNHEKNERINMVIQELGLDKVADSKVGT	180
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
Db	181	QFIRGVSGGERKRTSIAMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
Qy	241	SIHQPRYSIFKLFDSTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:             :             :	
Db	241	SIHQPRYSIFKLFDSTLLASGRLMFHGPAREALGYFASIGYNCEPYNNPADFFLDVING	300
Qy	301	DSTAVALNR-EEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:       :   :       : :       :           :	
Db	301	DSSAVVLSRADRDEGAQEPEEPPEKDTPLIDKLAIFYTNSSFFKDTKVELDQFSGGRKKK	360
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKND	419
		:   :     : :             :           : :         : :	
Db	361	KSSVYKEVTTYTTSFCHQLRWISRRSFKNLLGNPQASVAQIIVTIILGLVIGAIIFYDLKND	420
Qy	420	STGIQNRAGVLFFLTNTNOCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL	479
		:	
Db	421	PSGIQNRAGVLFFLTNTNOCFSSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSDLL	480
Qy	480	PMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL	539
		:         :     :	
Db	481	PMRMLPSIIFTCITYFLLGLKPAVGSSFFIMMFTLMMVAYSASSMALAIAAGQSVVSVATL	540
Qy	540	LMTICFVFMFIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFPCPLNATG	599
		:                   :     :	
Db	541	LMTISFVFMFIFSGLLVNLKTVVPWLSWLQYFSIPRYGFSALQYNEFLGQNFPCPLNVT	600

Qy 600 NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655  
|| |:| ||| ||| ||| || ||||:|||||||:||||||| |||||  
Db 601 NNTCSFAICTGAEYLENQGISLSAWGLWQNHVALACMMVIFLTIAYLKLLLLLKKYS 656

## RESULT 2

S19421

ATP-dependent permease ADP1 precursor - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein YCR011c; protein YCR105

C;Species: *Saccharomyces cerevisiae*

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C;Accession: S19421; S40914

R;Goffeau, A.; Purnelle, B.; Skala, J.

submitted to the Protein Sequence Database, March 1992

A;Reference number: S19420

A;Accession: S19421

A;Molecule type: DNA

A;Residues: 1-1049 &lt;GOF&gt;

A;Cross-references: UNIPROT:P25371; UNIPARC:UPI00001255FD; EMBL:X59720; NID:g1907116; PIDN:CAA42328.1; PID:g1907154; GSPDB:GN00003; MIPS:YCR011c

R;Purnelle, B.; Skala, J.; Goffeau, A.

Yeast 7, 867-872, 1991

A;Title: The product of the YCR105 gene located on the chromosome III from *Saccharomyces cerevisiae* presents homologies to ATP-dependent permeases.

A;Reference number: S40914; MUID:92160395; PMID:1789009

A;Accession: S40914

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-1049 &lt;PUR&gt;

A;Cross-references: UNIPARC:UPI00001255FD

R;Skala, J.; Purnelle, B.; Goffeau, A.

Yeast 8, 409-417, 1992

A;Title: The complete sequence of a 10.8 kb segment distal of *SUF2* on the right arm of chromosome III from *Saccharomyces cerevisiae* reveals seven open reading frames including the *RVS161*, *ADP1* and *PGK* genes.

A;Reference number: S25353; MUID:92327849; PMID:1626432

A;Contents: annotation

C;Genetics:

A;Gene: SGD:ADP1; MIPS:YCR011c

A;Cross-references: SGD:S0000604; MIPS:YCR011c

A;Map position: 3R

C;Superfamily: ATP-dependent permease ADP1; ATP-binding cassette homology

C;Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein

F;1-25/Domain: signal sequence #status predicted &lt;SIG&gt;

F;26-1049/Product: ATP-dependent permease ADP1 #status predicted &lt;MAT&gt;

F;26-324/Domain: extracellular #status predicted &lt;EXT&gt;

F;325-341/Domain: transmembrane #status predicted &lt;TM1&gt;

F;406-607/Domain: ATP-binding cassette homology &lt;ABC&gt;

F;423-430/Region: nucleotide-binding motif A (P-loop)

F;550-557/Region: nucleotide-binding motif B

F;794-810/Domain: transmembrane #status predicted &lt;TM2&gt;

F;829-845/Domain: transmembrane #status predicted &lt;TM3&gt;

F;878-894/Domain: transmembrane #status predicted &lt;TM4&gt;

F;909-925/Domain: transmembrane #status predicted &lt;TM5&gt;

F;938-954/Domain: transmembrane #status predicted &lt;TM6&gt;

F;1025-1041/Domain: transmembrane #status predicted &lt;TM7&gt;

F;50,114,165,221/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;429/Binding site: ATP (Lys) #status predicted

Query Match 24.9%; Score 835.5; DB 1; Length 1049;  
 Best Local Similarity 30.5%; Pred. No. 1.8e-50;  
 Matches 222; Conservative 134; Mismatches 257; Indels 115; Gaps 18;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
		:    : :   :     :             :	
Db	355	LGSSKSPIRLP-DEDAVNNFLQNEDDL-----ATLSFENITYSVPSINS-----DGVE	402
Qy	61	KEILSNINGIMKPG-LNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRP-ANFK	118
		: :   :   :   :   :   :   :   :   :   :   :	
Db	403	ETVLNEISGIVKPGQILAIMGGSGAGKTTLLDILAMKRKTGHVSGSIKVNIGISMDRKSFS	462
Qy	119	CNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKV	178
		:       : :         :         : :     :   :   :   :	
Db	463	KIIGFVDQDDFLLPTLTVFETVLNSALLRLPKALSFEAKKARVYKVLEELRIIDIKDRII	522
Qy	179	GTQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQ-GRT	237
		:     :       : :         :     :         :     :	
Db	523	GNEFDRGISGGEKRRVSIACELVTSPLVLFLDEPTSGLDASNANNVIECLVRLSSDYNRT	582
Qy	238	IIFSIIHQPRYSIFKLFDSTLLASGRIMFHGPAQEALGYFESAGYHCEAYNNPADFFLDI	297
		: :         :           : :   : :   : : :           : :	
Db	583	LVLSTHQPRSNIFYLFDKLVLLSKGEMVYSGNAKKVSEFLRNEGYICPDNYNIADYLIDI	642
Qy	298	-----INGDSTAV	305
		:	
Db	643	TFEAGPQGKRRRIRNISDLEAGTDTNDIDNTIHTTFTSSDGTQREWAHLAAHRDEIRS	702
Qy	306	ALNREEDFKATE---IIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQ-LSGGEKKKK	360
		:   :     :   :   :   :   :   :   :	
Db	703	LLRDEEDVEGTDGRRGATEIDLNTKLLHDK---YKDSVYYAELSQEIEEVLSEGDEESN	758
Qy	361	IT--VFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKN	418
		: : :     : :       :   : : :   :   :   :	
Db	759	VLNGDLPTGQQSAGFLQQLSILNSRSFKNMYRNPKLLGNLYLLTILLSLFLGTLYYNVSN	818
Qy	419	DSTGIQNRAGVLFLLTTNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDL	478
		:         :   :   :   :   :   :   :   :   :	
Db	819	DISGFQNRMGFFILTIFYGFVTFTGLSSFALERIIFIKERSNNYYSPLAYYISKIMSEV	878
Qy	479	LPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVAT	538
		:   : :   : :         :   : :   : :   : : :   : :	
Db	879	VPLRVVPPILLSLIVYPMTGLNMKDNAFFKICIGILILFNLGISLEILTIGIIFEDLNNSI	938
Qy	539	LLMTICFVFMIFSGLLV---NLTTIASWLSWLQYFSIPRYGFTALQHNEF-----	586
		:   : : :       :   :   :   :   :   :   :	
Db	939	ILSVLVLLGSLFLSGLFINTKNITNVA--FKYLKNFSVFYYAYESLLINEVKTLMLKERK	996
Qy	587	LGQNF-CPGLNATGNNPCNYATCTGEEYLVKQGI--DLSPWGLWKNHVALACMIVIFLTI	643
		:   :   :   :   :   :	
Db	997	YGLNIEVPG-----ATILSTFGFVVQNLVFDIK-----ILALFNVVFLIM	1036
Qy	644	AYLKLLFL	651
		: :	

Db 1037 GYLALKWI 1044

RESULT 3

FYFFW

white protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 31-Dec-1990 #sequence\_revision 17-Feb-1995 #text\_change 09-Jul-2004

C;Accession: S08635; S07263; S10240

R;Pepling, M.; Mount, S.M.

Nucleic Acids Res. 18, 1633, 1990

A;Title: Sequence of a cDNA from the Drosophila melanogaster white gene.

A;Reference number: S08635; MUID:90221897; PMID:2109311

A;Accession: S08635

A;Molecule type: mRNA

A;Residues: 1-687 <PEP>

A;Cross-references: UNIPROT:P10090; UNIPARC:UPI000011F0A1; EMBL:X51749; NID:g8825; PIDN:CAA36038.1; PID:g8826

R;O'Hare, K.; Murphy, C.; Levis, R.; Rubin, G.M.

J. Mol. Biol. 180, 437-455, 1984

A;Title: DNA sequence of the white locus of Drosophila melanogaster.

A;Reference number: S07263; MUID:85134865; PMID:6084717

A;Accession: S07263

A;Molecule type: DNA

A;Residues: 1-24, 'LIFEIPYHCRVTAD', 30-

334, 'ITLHLNSYPAWVPSVLPTTIRRTFTYRCWPLCPDGRSSPVIGSPRYG', 372-687 <OHA1>

A;Cross-references: UNIPARC:UPI000016B571; EMBL:X02974

A;Experimental source: strain Canton S

R;O'Hare, K.

submitted to the EMBL Data Library, June 1985

A;Reference number: S10240

A;Accession: S10240

A;Molecule type: DNA

A;Residues: 1-24, 'LIFEIPYHCRVTAD', 30-687 <OHA2>

A;Cross-references: UNIPARC:UPI000016BDFF; EMBL:X02974; NID:g10873; PIDN:CAA26716.1; PID:g10874

A;Experimental source: strain Canton S

C;Genetics:

A;Gene: white; w

A;Cross-references: FlyBase:FBgn0003996

A;Introns: 24/3; 116/1; 334/2; 439/3; 483/3

C;Superfamily: fruit fly white protein; ATP-binding cassette homology

C;Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein

F;113-317/Domain: ATP-binding cassette homology <ABC>

F;130-137/Region: nucleotide-binding motif A (P-loop)

F;261-265/Region: nucleotide-binding motif B

F;67,93,472,554,651/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.2%; Score 812; DB 1; Length 687;  
Best Local Similarity 32.1%; Pred. No. 4.4e-49;  
Matches 210; Conservative 134; Mismatches 251; Indels 60; Gaps 17;

Qy 5 NVEVFIPVSQGNTNGFPATASNDLKAFTGAVLSFHNICYRVKLKSGFLPCRKPVEKEIL 64  
|:::| |:| | : | | : :| | : | :|  
Db 74 NMDIFGAVNQ-----PGSGWRQLVNRTRGLFCNERHI-----PAPR---KHLL 113  
  
Qy 65 SNINGIMKPG-LNAILGPTGGGKSSLLDVLAARKDPSGL----SGDVLINGAPRPA-NFK 118



[http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7\\_142911\\_us-09-961-086a-1.rpr&ItemType=4&startByte=0](http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7_142911_us-09-961-086a-1.rpr&ItemType=4&startByte=0) (8 of 24)9/22/2008 12:04:28 PM



RESULT 5

C86441

probable ABC transporter [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: C86441

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C. Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L. J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C86441

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-646 <STO>

A;Cross-references: UNIPROT:Q9C6R7; UNIPARC:UPI00000AA9CB; GB:AE005172; NID:g11136734; PIDN: AAG31315.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

Query Match 23.1%; Score 774; DB 2; Length 646;  
Best Local Similarity 33.5%; Pred. No. 1.9e-46;  
Matches 213; Conservative 119; Mismatches 234; Indels 70; Gaps 22;

Qy	37	LSFHNICYRVKLK--SGFLPCRKPVEKEILSNINGIMKPG-LNAILGPTGGGKSSLLDVL	93
		:: :  : ::   :       :    :      :     ::	
Db	51	ITLKEVVYKVKIEQTSQCMGSWKSKEKTILNGITGMVCPGEFLAMLGPSGSGKTTLLSAL	110
Qy	94	AARKDPGLSGDVLINGAPRPANFKCNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMT	153
		:     :      :  :      :         :       ::	
Db	111	GGRLSKT-FSGKVMYNGQPFSGCIKRRTGFVAQDDVLYPHLTVWETLFFTALLRLPSSLT	169
Qy	154	NHEKNERINRVIQELGLDKVADSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLDEPT	213
		::       :: :  :    :    :        : : :	
Db	170	RDEKAEHVDRVIAELGLNRCTNSMIGGPLFRGISGGEKKRVSIGQEMLINPSLLLLDEPT	229
Qy	214	TGLDSSSTANAVLLLLKRMSKQGRTIIFSIIHQPRYSIFKLFDLTLASGRLMFHGPQEA	273
		:   :   : :: :   :    :: :       : :    :  :	
Db	230	SGLDSTTAHRIVTTIKRLASGGRTVVTTIHQPSSRIYHMFDKVLLSEGSPIIYYGAASSA	289
Qy	274	LGYFESAGYHCEAYNNPADFFLDIING--DSTAVALNREEDFKATEIIEPSKQDKPLIEK	331
		:       :       :      : :    :   :  :	
Db	290	VEYFSSLGFSTSLTVNPADLLLDLANGIPPDTQKETSEQEQTVKETL-VSAYEKNISTK	348

Qy 332 L-AEI-YVNSSFYKETKAELHQLSGGEKKKKITVFKEISYTTSFCHQL-----RWVSKRS 384  
| ||: | |: || | : |: |: | | : |  
Db 349 LKAELCNAESHSEYTKAAAKNL-----KSEQWCTTWYQFTVLLQRGVRERR 396

Qy 385 FKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKNDSTGIQNRAGVLFFLTNQCFSVS-S 443  
|:: : | |: | | : :: : |: | : || : | : :  
Db 397 FESF---NKLRIQVISVAFLG---GLLWW--HTPKSHIQDRTALLFFFSVFWGFYPLYN 448

Qy 444 AVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKA 503  
| | | ||:: | | || ||: |||: : : | ||: : | |: |: | |||  
Db 449 AVFTFPQEKRMILIKERSSSGMYRLSSYFMARNVGD-LPLELALPTAFVFIYWMGGLKPDP 507

Qy 504 DAFFVMMFTLMMVAYS---ASSMALAIAAGQSVVSVATLLMTICFVFMFIFSGLLVNLTT 560  
| : : |: | | : | | : | | | :: : : | | |  
Db 508 TTF---ILSLVLVLYSVLVAQGLGLAFGALLMNIKQATTLASVTTLVFLIAGGYVQ--Q 562

Qy 561 IASWLSWLQYFSIPRYGFTALQHNEFLGQNFPCPLNATGNNPCNYATCT-----GEEY 613  
| : : |: | | : | : | : | : | : | :  
Db 563 IPPFIVWLKYSYSYCYKLLL-----GIQYTDG---DYECSEKGVWCRVGDGFP 608

Qy 614 LVK-QGIDLSPWGLWKNHVALACMIVIFLTIAYLKL 648  
: | |: : | | : : |: | : : |: |  
Db 609 AIKSMGLN---NLWIDVFVMGVMLVGYRLMAYMAL 640

RESULT 6

D96553

hypothetical protein F5D21.6 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: D96553

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C. Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L. J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D96553

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-687 <STO>

A;Cross-references: UNIPROT:Q9C8K2; UNIPARC:UPI00000AB8E7; GB:AE005173; NID:g10092349; PIDN: AAG12758.1; GSPDB:GN00141

C;Genetics:

A;Gene: F5D21.6

A;Map position: 1

C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

Query Match		22.9%;	Score 767.5;	DB 2;	Length 687;
Best Local Similarity		32.6%;	Pred. No. 5.9e-46;		
Matches 212;		Conservative 118;	Mismatches 244;	Indels 77;	Gaps 21;
Qy	34	GAVLSFHNICYRVKLKSGFLPCRKPVEKEILSNINGIMKPG-LNAILGPTGGGKSSLLDV	92		
		::: : :       : :  :   :   :   ::     :			
Db	21	GAYLAWEDLTVVIPNFSGG-PTRR-----LLDGLNGHAEPGRIMAIMGPSGSGKSTLLDS	74		
Qy	93	LAARKDPSG-LSGDVLINGAPRPANFKCNSGYVVQDDVVMGTLTVRENLFSAALRLATT	151		
		: ::  ::     : :    ::     ::  :       ::			
Db	75	LAGRLARNVIMTGNLLLNKKARLDYGL-VAYVTQEDILMGTTLTVRETITYSAHLRLSSD	133		
Qy	152	MTNHEKNERINRVIQELGLDKVADSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLDE	211		
		:     : :          :             : :			
Db	134	LTKEEVNDIVEGTIIELGLQDCADRVIGNWHSRGVSGGERKRVSVALEILTRPQILFLDE	193		
Qy	212	PTTGLDSSSTANAVLLLLKRMSKQ-GRTIIFSIIHQPRYSIFKLFDSLTLASGRLMFHGPA	270		
		:    ::   :  : ::    ::       :          :    ::   :			
Db	194	PTSGLDSASAFFVIQALRNIARDGGRTVVSSIIHQPSSEVFALFDDLFLSSGETVYFGES	253		
Qy	271	QEALGYFESAGYHCEAYNNPADFFLDIINGDSTAVALNREEDFKATEIIEPSKQDKPLIE	330		
		:  : :    :     :          : :    :     :			
Db	254	KFAVEFFAEAGFPCPKKRNPSDHFLRCINSDFDTVTATLKGSQRIRE--TPATSD-PLMN	310		
Qy	331	-----KLAEIYVNSSFYKETKAELHQLSGGE-----KKKKITVFKEISYTTSFC	374		
		:       :    : : : :     :			
Db	311	LATSEIKARLVENYRRSVYAKSAKSRIRELASIEGHHGMEVRKGSEATWFK-----	361		
Qy	375	HQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKNDSTGIQNRAGVLEFLT	434		
		::     : : ::  : :   : : : :        :			
Db	362	-QLRTLTKRSFVNMCRDIGYYWSRIVYIYVVSFCVGTIFYDVGHSYTSILARVSCGGFIT	420		
Qy	435	TNQCFSSVSAVELFVVEKKLFIHEYISGYRVSYSYFLGKLLSDLLPMTMLPSIIFTCIVY	494		
		:  :    :    :         : :    : ::			
Db	421	GFMTFMSIGGFPSFIEEMKVIFYKERLSGYYGVSYYIISNYVSS-FPFLVAIALITGSITY	479		
Qy	495	FMLGLKPKAD--AFF-----VMMFTLMMVAYSASSMALAIAAGQSVVSVATLLMTI	543		
		: :      : :    : : :   :: : ::			
Db	480	NMVKFRPGVSHWAFFCLNIFFSVSVIESLMMVVASLVPNFLMGLITGAGIIGI--IMMTS	537		
Qy	544	CFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQ---HNEFLGQNFCPGLNATGN	600		
		:  :         :   :			
Db	538	GFF-----RLLPDLPKV----FWRYPISFMSYGSWAIQGAYKNDFLGLFEDPMF-----	582		
Qy	601	NPCNYATCTGEEYLK-QGIDLSP---WGLWKNHVALACMIVIFLTIAYLK	647		
		: :    : : :     :     ::  :			
Db	583	--AGEPKMTGEQVINKIFGVQVTHSKWWDLSAIVLILVCYRILFFIVLKLK	631		

RESULT 7  
A84509  
probable ABC transporter [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: A84509  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.;  
Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.;  
Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.  
M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.;  
Salzberg, S.L.; Fraser, C.M.; Venter, J.C.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: A84509  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-649 <STO>  
A;Cross-references: UNIPROT:Q9SIT6; UNIPARC:UPI00000A95A1; GB:AE002093; NID:g4558665; PIDN:  
AAD22683.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g13610  
A;Map position: 2  
C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-  
binding cassette homology

Query Match 22.3%; Score 746; DB 2; Length 649;  
Best Local Similarity 33.5%; Pred. No. 1.8e-44;  
Matches 203; Conservative 116; Mismatches 239; Indels 48; Gaps 18;

Qy	61	KEILSNINGIMKP-GLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAP-RPANFK	118
		:  :    :   :          :	
Db	60	KHVLKGVTCRAKPWEILAIIVGPGAGKSSLEILAAARLIPQ--TGSVYVKNRPVDRANFK	117
Qy	119	CNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKV	178
		:               :     : :       :   :	
Db	118	KISGYVTQKDTLFLPLTVEETLLFSAKLRL--KLPADELRSRVKSLVHELGLEAVATARV	175
Qy	179	GTQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSK-QGRT	237
		:   :        :       :         :   : :     : :	
Db	176	GDDSVRGISGGERRRVSIGVEVIHDPKVLILDEPTSGLDSTSALLIIDMLKXMAETRGRT	235
Qy	238	IIFSIIHQPRYSIFKLFDSLTLASGRIMFHGPAQEALGYFESAGYHCEAYNNPADFFLDI	297
		:       :       :       :       :     :   : :	
Db	236	IILTIHQPGFRIVKQFNSVLLLLANGSTLKGQSVQDLGVYLRNGLHPPLHENIVEFAIES	295
Qy	298	INGDSTAVALNREEDFKATEIIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEK	357
		:   :   : :       :           : :	
Db	296	I--ESITKQQRLQESRRAAHVLTLP---QTTLQEKRSQGES--KSGKFTLQQLFQQTR	348
Qy	358	KKKI-TVFKEISYTTSFCH----QLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAI	412
		:   : :   : : : :     : : : : :   :   :	
Db	349	VADVGTMNIAEFTRDFANSRLEETMILTHRFKSNIFRTKELFACRTVQMLGSGIVLGLI	408
Qy	413	YFGLKNDSTGIQNRAGVLFLLTTNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLG	472
		:     :   :   :   : :   : : :           : :	
Db	409	FHNLKDDLKGARERVGLFAFILTFLLTSTIEALPIFLQEREILMKETSSGSYRVSSYAVA	468
Qy	473	KLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAA---	529
		: : :   :     : :     : :   :   : : :	
Db	469	NGLV-YLPFLLILAILFSTPVYWLVLNPSFMAFLHFSLLIWLILYTANSVVVCFSALVP	527

Query Match 22.2%; Score 745; DB 2; Length 725;  
Best Local Similarity 29.3%; Pred. No. 2.4e-44;  
Matches 193; Conservative 130; Mismatches 274; Indels 62; Gaps 13;

http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7\_142911\_us-09-961-086a-1.rpr&ItemType=4&startByte=0 (13 of 24)9/22/2008 12:04:28 PM

```
Qy      273  ALGYFESAGYHCEAYNNPADFFLDII-----NGDSTAVALNREEDFKATEIIEPS--- 322
          |:|  |      |  ::| ||::      |      |  :      :|: :
Db      312  LPGFFSDFGRPIPEKENISEFALDLVRELEGSNEGTKALVDFNEKWQONKISLIQSAPQT 371

Qy      323  ---KQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKKITVFKEISYTTSFCHQLRW 379
          ||: |  |  |  :|:|  :: :|  :||  :      ::  ||      :
Db      372  NKLDQDRSLSLKEA---INASV---SRGKL--VSGSSRSNPTSMETVSSYANPSLFETFI 423

Qy      380  VSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKNDSTGIQNRAGVLFFLTNQC 439
          ::||  || :  |:      :|  :| | ::  :|: | :  | | |  :  |:      :
Db      424  LAKRYMKNWIRMPELVGTRIA TVMTGCLLATVYWKLDHTPRGAQERLTLFAFVVP TMEFY 483

Qy      440  SSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLPM TMLPSIIFTCIVYFMLGL 499
          :  |  :|: |: :|: |      || ||| :  |  ||  :  ||::|: |  :: :||
Db      484  CCLDNVPVFIQERYIFLRETTHNAYRTSSYVISHSLVS-LPQL LAPSLVFSAITFWTVGL 542

Qy      500  KPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLLMTICFV FMMIFSGLLVNLT 559
          :  |      :      :| ||:  |:      : :  ::      : :: ||  ||
Db      543  SGGLEGFVFYCLLIYASFWSGSSVVTFISGVVPNIMLCY MVSITYLAYCLLLSGFYVNRD 602

Qy      560  TIASWLSWLQYFSIPRYGFTALQHNEFLGQNFC-----PGLNATG----- 599
          |  : :|  | || :| : |:  |||  :  |      |:: :|
Db      603  RIPFYWTFWHYISILKYPYEAVLINEFDDPSRCFVRGVQVFDSTLLGGVSDSGKV KLET 662

Qy      600  -----NNPCNYATC--TGE EYLVKQGI-DLSPWGLWKNHVALACMIVIFLT IAYLKLLF 650
          :||  || :  | :|||  || |      |: :  | :  : ::||
Db      663  LSKSLRTKITESTCLRTGSDLLAQQGITQLSKWD-----CLWITFASGLFFRILF 712
```

RESULT 9

T45891  
ABC transporter-like protein - Arabidopsis thaliana  
N;Alternate names: protein F4P12.210  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T45891  
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23016  
A;Accession: T45891  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-739 <BLO>  
A;Cross-references: UNIPROT:Q9LFG8; UNIPARC:UPI00000A86EE; EMBL:AL132966  
A;Experimental source: cultivar Columbia; BAC clone F4P12  
C;Genetics:  
A;Map position: 3  
A;Note: F4P12.210  
C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

Query Match 22.1%; Score 741.5; DB 2; Length 739;  
Best Local Similarity 28.3%; Pred. No. 4.4e-44;  
Matches 193; Conservative 123; Mismatches 256; Indels 111; Gaps 17;

```
Qy      36  VLSFHNICYRVKLKSGF--LPC--RKPVE-----KEILSNINGIMKPG-LNAILGPT 82
          |||| :: | ||:|  |  ||      | :      | :|: |:|  :  | :  |:|| :
```

```

Db      87 VLSFKDLTYSVKIKKKFKPFPCCGNSPFDGNDMEMNNTKVLLNGISGEAREGEMMAVLGAS 146

Qy      83 GGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANF-KCNSGYVVQDDVVMGTLTVRENLQ 141
      | |||:|:| || | | ||: || :| | ||:| ||:| || | |
Db      147 GSGKSTLIDALANRISKESLRGDITLNGEVLESSLHKVISAYVMQDDLLFPMLTVEETLM 206

Qy      142 FSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGTQFIRGVSGGERKRTSIGMELI 201
      ||| || :::: :| |: :| :|| | : :| : ||||| ||:| ||| :||
Db      207 FSAEFRLPSSLSKSKKKKARVQALIDQLGLRNAAKTVIGDEGHRGVSGGERRRVSIGTDII 266

Qy      202 TDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIFSIIHQPRYSIFKLFDSTLLAS 261
      || ||||| ||:| |||:| |: :|:|:| | :| |||| | | | | |:
Db      267 HDPIILFLDEPTSGLDSTSAYMVVKVLQRIAQSGSIVIMSIHQPSYRILGLLDKLIFLSR 326

Qy      262 GRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIINGDSTAVALNREEDFKATEIIEP 321
      | :| | :| |: | :| ||:| | : :
Db      327 GNTVYSGSPTHLPQFFSEFGHPPIPENENKPEFALDLIR-----ELED 369

Qy      322 SKQDKPLIEKLAIEIYVNSSFYKETKAELHQLSGGEKKKKITVFKEI----- 367
      : | |:| |:|:|:| | | :| |:
Db      370 PEGTKSLVE-----FHKQWRAK--QTSSQSRRTNVSLKDAISASISRGKLVSGA 417

Qy      368 -----SYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKND 419
      :: | :: : || | | : :: :| |:| |:| |
Db      418 TNLRSSFQTFANPFWTEMLVIGKRSILNSRRQPELFGIRLGAVLVTGMILATIFWKLDNS 477

Qy      420 STGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL 479
      ||| | | | : :: |: :|:|:|:| | || || | : :
Db      478 PRGIQERLGFFAFAMSTTFYTCAEAIPVFLQERYIFMRETAYNAYRRSSYVLAHTIIS-I 536

Qy      480 PMTMLPSIIFTTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL 539
      | :: | | : :|| :: | ||:: :: || :: | | :
Db      537 PALIILSAAFAASTFSAVGLAGGSEGFLFFFFTILTAFWAGSSFTFLSGVVSHVMIGFT 596

Qy      540 LMTICFVFMFIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNEC--PGLNA 597
      :: : ::||| :: | : | | |:|:| : : ||| | |:
Db      597 VVVAILAYFLLFSGFFISRDRIPLYWIWFHYLSLVKYPYEGVLQNEFEDPTKCFVRGIQM 656

Qy      598 TGNNPCNYA-----TC--TGEEYLVKQGI-DLSPWG-LWKNH 630
      |:| || || :| :||| :| | |
Db      657 FDNSPLGQVPTAVKISLLKSMGVLGINVTAETCVTTGIDILKQQGITEISKWNCLW--- 713

Qy      631 VALACMIVIFLTIA---YLKLLF 650
      :|:| : :||
Db      714 -----ITVAWGFFFRVLF 726

```

## RESULT 10

H96552

hypothetical protein F5D21.8 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: H96552

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.

Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.;  
Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.;  
Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti,  
R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.;  
Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.  
J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.;  
Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: H96552  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-678 <STO>  
A;Cross-references: UNIPROT:Q9C8J8; UNIPARC:UPI00000A045D; GB:AE005173; NID:g10092361; PIDN:  
AAG12770.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F5D21.8  
A;Map position: 1  
C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-  
binding cassette homology

Query Match 22.0%; Score 739; DB 2; Length 678;  
Best Local Similarity 30.3%; Pred. No. 5.8e-44;  
Matches 186; Conservative 137; Mismatches 248; Indels 42; Gaps 13;

Qy 61 KEILSNINGIMKPG-LNAILGPTGGGKSSLLDVLAARKDPSG---LSGDVLINGAPRPAN 116  
| :|: :|| :| : ||:|:| |||:| | | :| :|| ||:| | :  
Db 28 KRLNGVNGCGEPNRILAIMGPSGSGKSTLLDALAGRL--AGNVMSGKVLVNGKKRRLD 85  
  
Qy 117 FKCNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADS 176  
| : || |:|:|:| | | | : || ||| : :| | : : | :|:|:| :|  
Db 86 FGA-AAYVTQEDVLLGTLTVRESISYSAHLRLPSKLTREEISDIVEATITDMGLEECSDR 144  
  
Qy 177 KVGTFQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGR 236  
:| :|:|:| | | | :|:| | | | | | | :| | : :| :| :  
Db 145 TIGNWHLRGISGGEKKRLSIALEVLTGPSLLFLDEPTSGLDSASAFFVVQILRNIASSGK 204  
  
Qy 237 TIIFSIIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLD 296  
|:| | | | :| | | | | :| :| :| :| :| :| :| :| :|  
Db 205 TVVSSIIHQPSGEVFALFDDLLLLSGGETVYFGEAESATKFFGEAGFPCPSRRNPSDHFLR 264  
  
Qy 297 IINGD----STAVALNR---EEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAEL 349  
:| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 265 CVNSDFDNVTAALVESRRINDSSFSHLQHETNTLDPL-DDIPTAEIRTTLVKRFKCSL 323  
  
Qy 350 HQLSGGEKKKKITVF-----KEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTV 403  
: : : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 324 YAAASRARIQEIASIVGIVTERKKGSQTNWWKQLRILTQRSFINMSRDLGYYWMRIAVYI 383  
  
Qy 404 VLGLVIGAIYFGLKNDSTGIQNRAGVLFLLTTNQCFSSVSAVELFVVEKKLFIHEYISGY 463  
|| : :|:|:| : :| : :| :| :| :| :| :| :| :| :| :| :| :|  
Db 384 VLSICVGSIFFNVGRNHTNVMSTAACGGFMAGFMFMSIGGFQSFIEEMKVFSRERLNGH 443  
  
Qy 464 YRVSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSM 523



```

      | | : | : | | | | | : | : : | : | : : | | | : | : |
Db      444 YGVAVYTVSNLLSS-LPFIILMCLSTSSITIYMVRFQSGGSHFFYNCLDLICAITTVESC 502

Qy      524 ALAIAAGQSVSVSATLLMTIC----FVFMIMFSGLLVNLTTIASWLSWLQYFSIPRYGFT 579
      : | | : | | | : : : | | : | | |
Db      503 MMMIAS-----VVPNFLMGVMLGAGYIGIMVLSAGFFRFFPDLPVFWRYPVSYINYGAW 557

Qy      580 ALQ---HNEFLGQNFCPGLNATGNNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACM 636
      | | | | | : | : : : | | : : : | | : : |
Db      558 ALQGAYKNEMIGVEY-----DSPLPLVPKMKGELILQTVLGINPESSKWLDLAVVMM 609

Qy      637 IVIFLTIAYLKLL 649
      | : | | | : : |
Db      610 ILIGYRIAFFAIL 622
```

RESULT 11

T47650  
ABC transporter-like protein - Arabidopsis thaliana  
N;Alternate names: protein T15C9.110  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T47650  
R;Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z24470  
A;Accession: T47650  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-708 <MEW>  
A;Cross-references: UNIPROT:Q9M2V5; UNIPARC:UPI00000A627B; EMBL:AL132970  
A;Experimental source: cultivar Columbia; BAC clone T15C9  
C;Genetics:  
A;Map position: 3  
A;Note: T15C9.110  
C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

Query Match 21.7%; Score 726.5; DB 2; Length 708;  
Best Local Similarity 30.1%; Pred. No. 4.7e-43;  
Matches 198; Conservative 123; Mismatches 282; Indels 55; Gaps 15;

```

Qy      36 VLSFHNICYRVKLKSGFLPCRKPVE--KEILSNINGIMKPG-LNAILGPTGGGKSSLLDV 92
      : | | : | | | : | : | : | : | : | : | : | : |
Db      62 LLSFNNLSYNVVLRRRFDfsRRKTASVKTLDDITGEARDGEILAVLGGSGAGKSTLIDA 121

Qy      93 LAARKDPSGLSGDVLINGAP--RPANFKCNSGYVVQDDVVMGTLTVRENLQFSAALRLAT 150
      | | | | | : | | : | | | : | | : | : |
Db      122 LAGRVAEDSLKGTVTLNGEKVLQSRLKVISAYVMQDDLLFPMLTVKETLMFASEFRLPR 181

Qy      151 TMTNHEKNERINRVIQELGLDKVADSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLD 210
      : : : | | : : | : | | | | : : : | | : | | |
Db      182 SLPKSKKMervETLIDQLGLRNAADTVIGDEGHRGVSGGERRRVSIGIDIHDPILLFLD 241

Qy      211 EPTTGLDSSTANAVLLLLKRMSKQGRTIIFSihQPRYSIFKLFDsLTLLASGRLMFHGPA 270
      | | : | | : | : | : | : | | | | | : | : | : |
Db      242 EPTSGLDSTNAFMVVQVLKRIAQSGSVVIMSIHQPSARIIGLLDRLIILSHGKSVFENGSP 301
```

```
Qy      271 QEALGYFESAGYHCEAYNNPADFFLDII-----NGDSTAVALNREEDFKATEIIEPSKQD 325
          :| | |           | :| ||:|           : : | : | :: : : |
Db      302 VSLPSFFSSFGRIPEKENITEFALDVIRELEGSSEGTRDLVEFNEKWQQNQATARATTQS 361

Qy      326 KPLIEKLAIEIYVNSSFYKETKAELHQLSGGEKKKKITVFKEISYTTSFCHQLRWVSKRSF 385
          : :: : |           :: :| :||           |:: ||           : ::||
Db      362 RVSLKEAIAASV-----SRGKL--VSGSSGANPISMETVSSYANPPLAETFILAKRYI 412

Qy      386 KNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKNDSTGIQNRAGVLFFLTNNQCFSSVSAV 445
          || : | : :| :| ||:: :|: | | | | | | : : :
Db      413 KNWIRTPELIGMRIGTVMVTGLLLATVYWRLDNTPRGAQERMGFFAFGMSTMFYCCADNI 472

Qy      446 ELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADA 505
          :|: |: |: |: | || ||| : | || : || | :: :|| ::
Db      473 PVFIQERYIFLRETTNAYRTSSYVISHALVS-LPQLLALSIAFAATTFTWTVGLSGGLES 531

Qy      506 FFVMMFTLMMVAYSASSMALAIAAGQSVSVATLLMTICFV-FMMIFSGLLVNLTTIASW 564
          || : : :| ||: | :| : : :|| :: : : : | :| | :
Db      532 FFYYCLIIYAAFWSGSSIIVTFI-SGLIPNVMSYMTIAYLSYCLLLGGFYINRDRIPLY 590

Qy      565 LSWLQYFSIPRYGFTALQHNEFLGQNFC-----PGLNATG 599
          | | | :| : |: ||| : | | : :
Db      591 WIWFHYISLLKYPYEAVLINEFDDPSRCFVKGVQVFDGTLLEVSHVMKVKLLDTLSGSL 650

Qy      600 NNPCNYATC--TGEEYLVKQGI-DLSPWG-LWKNHVALACMIVIFLTIAYLKLLFLKK 653
          :|| || : |:||| || | || : || : | : || ||| |
Db      651 GTKITESTCLRTGPDLLMQQGITQLSKWDCLW---ITLAWGL-FFRILFYLSLLFGSK 704
```

RESULT 12

T08934

hypothetical protein F27G19.20 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: T08934

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.

submitted to the Protein Sequence Database, May 1999

A;Reference number: Z16519

A;Accession: T08934

A;Molecule type: DNA

A;Residues: 1-635 <BEV>

A;Cross-references: UNIPROT:Q9SZR9; UNIPARC:UPI00000A1E91; EMBL:AL078467; GSPDB:GN00062; ATSP:F27G19.20

A;Experimental source: cultivar Columbia; BAC clone F27G19

C;Genetics:

A;Gene: ATSP:F27G19.20

A;Map position: 4

A;Introns: 38/3; 253/1; 304/1; 414/3

C;Superfamily: fruit fly white protein; ATP-binding cassette homology

```
Query Match          21.5%;  Score 721;  DB 2;  Length 635;
Best Local Similarity 30.0%;  Pred. No. 9.7e-43;
Matches 203;  Conservative 121;  Mismatches 284;  Indels 68;  Gaps 18;
```

```
Qy      1  MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSG---FLPCRK 57
```

Db 1 MDNQEVSMDEVETPIAKTNDRLPFSIFKKANNPVTLKFENLVYTVKLKDSQGCFCGKNDK 60

Qy 58 PVEKEILSNINGIMKPG-LNAILGPTGGGKSSLLDVLAAR--KDPGLSGDVLINGAPRP 114

Db 61 TEERTILKGLTGIVKPGEILAMLGPSGSGKTSLLTALGGRVGEKGKLTGNISYNNKPLS 120

Qy 115 ANFKCNSGYVVQDDVVMGTLTVRENLFSAALRLATTMTNHEKNERINRVIQELGLDKVA 174

Db 121 KAVKRTTGFVTQDDALYPNLTVTETLVFTALLRLPNSFKKQEKIKQAKAVMTELGLDRCK 180

Qy 175 DSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQ 234

Db 181 DTIIGGPFRLRGVSGGERKRVSIGQEILINPSLLFLDEPTSGLDSTTAQRIVSILWELARG 240

Qy 235 GRTIIFSIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYN-NPADF 293

Db 241 GRTVVTTIHQP-----SKGNPVYFGLGSNAMDYFASVGYSPLVERINPSDF 286

Qy 294 FLDIINGDSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYK-----E 344

Db 287 LLDIANGKPLL-----ISCWPSVGSDESQRPEAMKAALV----AFYKTNLLDSVINE 335

Qy 345 TKAELHQLSGGEKKKKITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQA--SIAQIIVT 402

Db 336 VKGQDDLCKNPRESSRVATNTYGDWPTTWWQQFCVLLKRGLKQRRHDSFSGMKVAQIF-- 393

Qy 403 VVLGLVIGAIYFGLKNDSTGIQNRAGVLFFLTNTQCFSSV-SAVELFVVEKKLFIHEYIS 461

Db 394 -IVSFLCGLLWWQTK--ISRLQDQIGLLFFISSFWAFFPLFQQIFTFPQERAMLQKERSS 450

Qy 462 GYYRVSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSAS 521

Db 451 GMYRLSPYFLSRVVGDLPMELILPTCFLVITYWMAGLNHNLANFFVTLLVLLVHVLVSG 509

Qy 522 SMALAIAGQSVSVATLLMTICFVFMIFSGLLVNLTTIASWLSWLQYFSIPRYGETAL 581

Db 510 GLGLALGALVMDQKSATTLGSMVIMLTFLLAGGYVQHVPV--FISWIKYVSIGYYTYKLL 567

Qy 582 QHNEFLGQNFPCPLNATGNN---PCNYATCTGEEYL-VKQGIDLSPWGLWKNHVALACMI 637

Db 568 ----ILGQYTANELYPCGDNGKLRCHVGDFEGIKHIGFNSGL-----VSALALTAML 615

Qy 638 VIFLTIAYLKLLFLKK 653

Db 616 VVYRVIAIALTRIGK 631

RESULT 13

G84791

probable ABC transporter [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: G84791

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.;

Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H. M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.  
Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84791

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-755 <STO>

A;Cross-references: UNIPROT:Q9ZUT0; UNIPARC:UPI00000A1BA9; GB:AE002093; NID:g4056489; PIDN: AAC98055.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g37360

A;Map position: 2

C;Superfamily: *Arabidopsis thaliana* probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

Query Match 21.5%; Score 720.5; DB 2; Length 755;  
Best Local Similarity 28.3%; Pred. No. 1.3e-42;  
Matches 202; Conservative 131; Mismatches 282; Indels 99; Gaps 21;

```

Qy      2 SSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGA-VLSFHNICYRVKLKSGFLP---CRK 57
      || : : |:: : | : || : : |||| :: | ||:: | | ||:
Db     63 SSRALGIASPINSA-ASSFNWASAPASSISSSPFVLSFTDLTYSVKIQKKFNPLACCRR 121

Qy     58 PVE-----KEILSNINGIMKPG-LNAILGPTGGGKSSLLDVLAARKDPSSLGSDVLING 110
      | :: |:: : | : |:: | : |||:: | | | | : : ||
Db    122 SGNDSSVNTKILLNGISGEAREGEMMAVLGASGSGKSTLIDALANRIAKDSLRSITLNG 181

Qy    111 APRPANF-KCNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELG 169
      :: | | ||::||:: || | | || | | :: : | : : ||
Db    182 EVLESSMQKVISAYVMQDDLLFPMLTVEETLMFSAEFLRPSLSKKKKKARVQALIDQLG 241

Qy    170 LDKVADSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLLK 229
      | | : : | : |||||:: | || : | | |||||:: | ||:: | : : |
Db    242 LRSAAKTVIGDEGHRGVSGGERRRVSIGNDIIHDPIILFLDEPTSGLDSTSAYMVIKVLQ 301

Qy    230 RMSKQGRTIIFSIIHQPRYSIFKLFDSTLLASGRMLFMHGPQEALGYFESAGYHCEAYNN 289
      |::: | : | |||| | | | | : | : : | : | : |
Db    302 RIAQSGSIVIMSIHQPSYRIMGLLDQLIFLSKGNVTYSGSPTHLPQFFSEFKHPIPENEN 361

Qy    290 PADFFLDIINGDSTAVALNREEDFKATEIIIEPSKQDKPLIE-----KLAEIYVN---- 338
      : | ||: | || :: : : |||: | | | |
Db    362 KTEFALDLI-----RELEYS-----TEGTKPLVEFHKQWRAKQAPSYNNNNKR 404

Qy    339 ----SSFYKETKAELHQ---LSGGEKKKKITVFKEI-SYTTSFCHQLRWVSKRSFKNLLG 390
      || : | : : : || : : | : : : ||: |
Db    405 NTNSSLKEAITASISRGKLVSGATNNNSSLNTPSFQTFANPFWIEMIVIGKRAILNSRR 464

Qy    391 NPQASIAQIIIVTVVLGLVIGAIYFGLKNDSTGIQNRAGVLFLLTTNQCFSSVSAVELFVV 450
      | : : : : | :: : : | | | | | : : : | : : |
Db    465 QPELLGMRLGAVMVTGIILATMFTNLDNSPKGAQERLGFFAFAMSTTFYTCAEAIPVFLQ 524

Qy    451 EKKLFIHEYISGYRVSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMM 510
      | : : | | || | : : : | : : | | : : || | |
Db    525 ERYIFMRETAYNAYRRSSYVLSQSIIS-IPALIVLSASFAATTFWAVGLDGGANGFFFFY 583

```

Qy 511 FTLMMVAYSASSMALAIAAGQSVVSVATLLMTICFVFMIFSGLLVNLTTIASWLSWLQY 570  
||:: ::|| :: | : :: : ::|| :: | : | |  
Db 584 FTILASFWAGSSFVTFLSGVIPNVMLGFTVVVAILAYFLLFSGFFISRDRIPVYWLWFHY 643  
  
Qy 571 FSIPRYGFTALQHNEFLGQN----FCPGLNATGNNP-----C 603  
| : :| : : ||| || | | : | :|  
Db 644 ISLVKYPYEGVLQNEF--QNPTRCFARGVQLFDNSPLGEFPNDVKVNLLKSMGVLGTNV 701  
  
Qy 604 NYATC--TGEEYLVKQGI-DLSPWG-LWKNHVALACMIVIFLTIA---YLKLLF 650  
|| || : | :||| | :| || :| : :||  
Db 702 TAETCVTTGIDILKQQGITDISKWNCLW-----ITVAWGFFFRVLF 742

RESULT 14  
T02567

probable ATP-binding cassette protein T16B24.1 - Arabidopsis thaliana  
N;Alternate names: protein F12L6.1  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T02567; T00545; C84816  
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S. M.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C. submitted to the EMBL Data Library, August 1998  
A;Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.  
A;Reference number: Z14679  
A;Accession: T02567  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-740 <ROU>  
A;Cross-references: UNIPROT:O80946; UNIPARC:UPI00000AC233; EMBL:AC004697; NID:g3402671; PIDN:AAC28975.1; PID:g3402672  
A;Experimental source: cultivar Columbia  
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C. submitted to the EMBL Data Library, July 1998  
A;Description: Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.  
A;Reference number: Z14168  
A;Accession: T00545  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-362 <ROW>  
A;Cross-references: UNIPARC:UPI00001746D9; EMBL:AC004218; NID:g3355463; PIDN:AAC27826.1; PID:g3355464  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H. M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C. Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: C84816  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-740 <STO>

A;Cross-references: UNIPARC:UPI00000AC233; GB:AE002093; NID:g3402672; PIDN:AAC28975.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g39350; T16B24.1; F12L6.1

A;Map position: 2

C;Superfamily: Arabidopsis thaliana probable ATP-binding cassette protein F12L6.1; ATP-binding cassette homology

C;Keywords: ATP

F;110-310/Domain: ATP-binding cassette homology <ABC>

Query Match 21.4%; Score 717.5; DB 1; Length 740;  
Best Local Similarity 28.0%; Pred. No. 2.1e-42;  
Matches 193; Conservative 120; Mismatches 294; Indels 83; Gaps 16;

Qy	26	NDLKAFTEGAVLSFHNICYRVKLK-----SGFLPCR-----KPVEKEILSNING	69
		:     ::     :   :   :	
Db	56	NDGYMRTVPFVLSFDNLTYNVSVRPKLDFRNLFPRRTEDPEIAQTARPKTKTLLNNISG	115
Qy	70	IMKPG-LNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPAN-FKCNSGYVVQD	127
		:   :   :   :       :       :   :       :	
Db	116	ETRDGEIMAVLGASGSGKSTLIDALANRIAKGSLKGTVKLNGETLQSRMLKVISAYVMQD	175
Qy	128	DVVMGTLTVRENLFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGTQFIRGVS	187
		: :           :     : :   :   :   :   :	
Db	176	DLLFPMLTVEETLMFAAEFRLPRSLPKSKKKLRVQALIDQLGIRNAAKTIIGDEGHRGIS	235
Qy	188	GGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIFSIIHQPRY	247
		:       : :     :         :     : :     :         :	
Db	236	GGERRRVSIGIDIHDPILLFLDEPTSGLDSTSAFMVVVKVLKRIAQSGSIVIMSIIHQPSH	295
Qy	248	SIFKLFDSLTLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDII-----NG	300
		:         :   : :   :   :   :     :	
Db	296	RVLGLLDRLIFLSRGHTVYSGSPASLPRFFTEFGSPIPENENRTEFALDLIRELEGSAGG	355
Qy	301	DSTAVALNREEDFKATEIIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQ---LSGGEK	357
		:   :     : :   :   :   :   : :   : :	
Db	356	TRGLIEFNK----KWQEMKKQSNRQPPLTPP-SSYPNLTKEAIAASISRGKLVSGGES	410
Qy	358	-----KKKITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGA	411
		: : : :   : :   : :   : :   : :   : :	
Db	411	VAHGGATTNTTTLAVPAFANPMWIEIKTLSKRSMNLNSRRQPELFGIRIASVVITGFILAT	470
Qy	412	IYFGLKNDSTGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFL	471
		: : :       :       : : :   : : :   : : :	
Db	471	VFWRLDNSPKGVQERLGFFAFAMSTMFYTCADALPVFLQERYIFMRETAYNAYRRSSYVL	530
Qy	472	GKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQ	531
		:   :   :   : :   : : :   : :   : :	
Db	531	SHAIVS-FPSLIFLSVAFAATTYWAVGLDGGLTGLLFYCLIILASFWSGSSFVTFSLGVV	589
Qy	532	SVVSVATLLMTICFVFMIMFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNF	591
		: : : : :   :   :   : :   : :   : :	
Db	590	PSVMLGYTIVVAILAYFLLFSGFFINRNRIPDYWIWFHYMSLVKYPYEAVLQNEFSDATK	649
Qy	592	C--PGLNATGNNP-----CNYATC--TGEEYLVKQG-IDLSP	623
		:     :     :   :   :	

Db 650 CFVRGVQIFDNTPLGELPEVMKLLKLLGTVSKSLGVTISSTTCLTTGSDILRQQGVVQLSK 709

Qy 624 WGLWKNHVALACMIVIFLTIA---YLKLLF 650

| :|:|:| : :||

Db 710 WN-----CLFITVAFGFFFRILF 727

RESULT 15

E96742

probable ABC transporter F17M19.11 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: E96742

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C. Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L. J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96742

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-609 <STO>

A;Cross-references: UNIPROT:Q9C8W6; UNIPARC:UPI000009EF81; GB:AE005173; NID:g6978921; PIDN: AAF34313.1; GSPDB:GN00141

C;Genetics:

A;Gene: F17M19.11

A;Map position: 1

C;Superfamily: fruit fly white protein; ATP-binding cassette homology

Query Match 21.3%; Score 712.5; DB 2; Length 609;

Best Local Similarity 31.0%; Pred. No. 3.6e-42;

Matches 190; Conservative 125; Mismatches 256; Indels 41; Gaps 16;

Qy 56 RKPVEKEILSNINGIMKPG-LNAILGPTGGGKSSLLDVLAAARKDPSGLSGDVLLINGAPRP 114

| | : ||| : |:: || |:|::| |||::: :| | | |:| :|||

Db 23 RSTEERTILSGVTGMISPGEFMAVLGPGSGKSTLLNAVAGRLHGSNLTGKILINDGKIT 82

Qy 115 ANFKCNSGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVA 174

:|:| |||:: ||||| | | | ||| :| | | || |||||

Db 83 KQTLKRTGFVAQDDLLYPHLTVRETLVFVALLRLPRSLTRDVKLRRAESVISELGLTKCE 142

Qy 175 DSKVGTQFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLLKRMsk- 233

:: || ||||:|||||| || ||: :|:| | ||||:||||: | :: | ::

Db 143 NTVVGNTFIRGISGGERKRVsIAHELLINPSLLVLDEPTSGLDATAALRLVQTLAGLAHG 202

Qy 234 QGRTIIFSihQPRYSIFKLFDsLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADF 293

	: : ::       : : : : :   :  : :    : : :       :	
Db	203 KGKTVVTSIHQPSSRVFQMFDTVLLLSEGGKCLFVGKGRDAMAYFESVGFSPAFFPMNPADF	262
Qy	294 FLDIING--DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEI----YVNSSFYKETKA	347
	:    :   : : :   :   :   :   :	
Db	263 LLDLANGVCQTDGVTEREKPNVRQTLVTAYDTLLAPQVKTCIEVSHFPQDNARFVKT---	319
Qy	348 ELHQLSGGEEKKKKITVFKEISYTTSFCHQL-RWVSKRSFKNLLGNPQASIAQIIVTVVLG	406
	: :    : : :       : :   : :	
Db	320 ---RVNNGGGITTICIA-----TWFSQLCILLHRLKERRHESEF-----DLLRIFQVVAAS	365
Qy	407 LVIGAIYFGLKNDSTGIQNAGVLFLLTTN-QCFSSVSAVELFVVEKKLFIHEYISGYR	465
	:  : : :  : :   : : :  :      : :	
Db	366 ILCGLMWW--HSDYRDVHDLGLLFFISIFWGVLPSPFNAVFTFPQERAIFTRERASGMYT	423
Qy	466 VSSYFLGKLLSDLLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMAL	525
	: : : : :    : :     : : :    : :  :  : : :	
Db	424 LSSYFMAHVLGSLSMELVLPASFLT-FTYWMVYLRPGIVPFLTLVLLLYVLASQGLGL	482
Qy	526 AIAAGQSVSVSATLLMTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGF---TALQ	582
	:   : : :  : : :     :   :  :      :  :	
Db	483 ALGAAIMDAKKASTIVTVTMLAFVLTGGYYVN--KVPSGMVWMKYVSTTFYCYRLLVAIQ	540
Qy	583 HNEFLGQNFCEP--GLNATGNNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIF	640
	:   :   : :   :      : : :    :  :     :	
Db	541 YGS--GEEILRMLGCDSKGKGASAAATSAGCRFVEEEVI--GDVGMWTSVGVLFLMFFGY	596
Qy	641 LTIAYLKLLFLK	652
	: :    :	
Db	597 RVLAYLALRRIK	608

Search completed: September 18, 2008, 22:09:13  
Job time : 43 secs